Simulation of Decision Making Process Using Quantum Parallelism

Michail Zak
Ultracomputing Group Mail Stop 126-347
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Abstract

A quantum device simulating human decision making process is introduced. It consists of quantum recurrent nets generating stochastic processes which represent the motor dynamics, and of classical neural nets describing evolution of probabilities of these processes which represent the mental dynamics. The autonomy of the decision making process is achieved by a feedback from mental to motor dynamics which changes the stochastic matrix based upon the probability distributions. This feedback replaces an unavailable external information by an internal knowledgebase stored in the mental model in the form of probability distributions. As a result, the coupled motor-mental dynamics is described by a nonlinear version of Markov chains which can decrease entropy without an external source of information. Applications to common sense based decisions are discussed.

1. Introduction

A human common sense has always been a mystery for physicists, and an obstacle for artificial intelligence. It was well understood that human behavior, and in particular, the decision making process, is governed by feedbacks from the external world, and this part of the problem was successfully simulated in the most sophisticated way by control systems. However, in addition to that, when the external world does not provide sufficient information, a human turns for "advise" to his experience, and that is associated with a common sense. In this paper, by common sense we will understand a feedback from the self-image (a concept adapted from psychology), and based upon that, we will propose a physical model of common sense in connection with the decision making process.

A decision making process can be modeled by a time evolution of a vector π whose components $\pi_i(i = 1,2\ldots N)$ represent a probability distribution over N different choices. The evolution of this vector can be written in the form of a Markov chain:

$$\pi_i(t+\tau) = \pi_i(t)P, \quad \sum_{i=1}^{N} \pi_i = 1, \ \sum_{j=1}^{N} p_{ij} = 1, \qquad 0 \le \pi_i \le 1, \quad 0 \le p_{ij} < 1 \tag{1}$$

where P is the transition matrix representing a decision making policy. If $P = Const$, the process (1) approaches some final distribution π^{∞} regardless of the initial state π^{o}. In particular, in the case of doubly-stochastic transition matrix, i.e., when

$$\sum_{j=1}^{N} p_{ij} = 1, \quad and \quad \sum_{i=1}^{N} p_{ij} = 1 \tag{2}$$

all the final choices become equally probable:

$$\pi_i = \pi_j = 1/N \tag{3}$$

i.e., the system approaches its thermodynamics limit which is characterized by the maximum entropy. When the external world is changing, such a rigid behavior is unsatisfactory, and the matrix P has to be changed accordingly, i.e., $P = P(t)$. Obviously this change can be implemented only if the external information is available, and there are certain sets of rules for correct responses. However, in real world situations, the number of rules grows exponentially with the dimensionalities of external factors, and therefore, any man-made device fails to implement such rules in full.

The main departure from this strategy can be observed in human approach to decision making process. Indeed, faced with an uncertainty, a human uses a "common sense" approach based upon his previous experience and knowledge in the form of certain invariants or patterns of behavior which are suitable for the whole class of similar situations. Such an ability follows from the fact that a human possesses a self-image, and interacts with it. This concept which is widely exploited in psychology has been known as far back as to ancient philosophers, but so far its mathematical formalization has never been linked to the decision making model (1).

First we will start with an abstract mathematical question: can the system (1) change its evolution, and consequently, its limit distribution, without any external "forces"? The formal answer is definitely positive. Indeed, if the transition matrix depends upon the current probability distribution

$$P = P(\pi) \tag{4}$$

then the evolution (1) becomes nonlinear, and it may have many different scenarios depending upon the initial state π^o. In particular case (2), it can "overcome" the second law of thermodynamics decreasing its final entropy by using only the "internal" resources. The last conclusion illuminates the Schrödinger statement [2] that 'life is to create order in the disordered environment against the second law of thermodynamics." Obviously this statement cannot be taken literally — as will be shown below, eq. (1) subject to the condition (4) describes the system which is not isolated, and therefore, the result stated above does not violate the second law of thermodynamics. In order to discuss the physical meaning of the condition (4), let us turn to Eq. (1) and introduce the underlying stochastic process. The latter can be simulated by a quantum device represented by quantum recurrent nets (QRN) [3] , and we will start with a brief description of that device.

The simplest QRN is described by the following set of difference equations with constant time delay τ

$$a_i(t+\tau) = \sigma_1\left\{\sum u_{ij}(t)a_j(t)\right\}, i.e., \{a_o a_1 \ldots a_N\} \to \{0,0\ldots \underset{\uparrow_i}{1} \ldots 00\} \tag{5}$$
$$i = 1,2\ldots N$$

where a_j is the input to the network at time t, u_{ij} is a unitary operator defined by the corresponding Hamiltonian of the quantum system, and σ_1 is a measurement operator (in the computational basis) that has the effect of projecting the evolved state into one of the eigenvectors of σ_1. The curly brackets are intended to emphasize that σ_1 is to be taken as a measurement operation with the effect similar to those of a sigmoid function in classical neural networks. Obviously, the outputs $a_i(t+\tau)$ are random because of the probabilistic nature of quantum measurements. As shown in [3], these outputs form a Markovian stochastic process with the probabilities evolving according to the chain (1) and

$$p_{ij} = |u_{ji}|^2, \quad \sum_{j=1}^{N} p_{ij} = 1, \quad \sum_{i=1}^{N} p_{ij} = 1, \quad p_{ij} \geq 0, \quad i,j = 1,2,\ldots N \tag{6}$$

is the NxN doubly-stochastic matrix which is uniquely defined by the unitary matrix U. Each element of this matrix represents the probability that the i[th] eigenvector as an input produces j[th] eigenvector as an output:

$$\{00 \quad 0\underset{\underset{i}{\uparrow}}{1}0 \quad 0\} \rightarrow \{00 \quad 0\underset{\underset{j}{\uparrow}}{1}0 \quad 0\} \tag{7}$$

In a special case when

$$p_{ij} > 0; \quad i,j = 1,2,\ldots N$$

the Markov process is ergodic, i.e., the solution to Eq. (1) approaches an attractor (3)

which is unique and it does not depend upon the initial value π_0 at t=0. Only this case will be considered in this paper. Thus, Eq. (5) describes the evolution of the vector

$$\{a_1 \ldots a_n\} = <\varphi|, \quad \sum_{j=1}^{N} a_j^2 = 1 \tag{8}$$

representing a quantum state in a Hilbert space, and all the components $(a_{i,} u_{ij})$ are to be actually implemented . This evolution is irreversible, nonlinear and nondeterministic because it includes measurement operations.

On the other hand, the vector

$$(\pi_1, \pi_2 \ldots \pi_n) = \pi, \quad \sum_{j=1}^{n} \pi_j = 1, \quad \pi_i > 0, \tag{9}$$

as well as the stochastic matrix p_{ij} exist only in an abstract Euclidean space: they never appear explicitly in physical space. The evolution (1) is also irreversible, but unlike (5), it is linear and deterministic.

So far we have simulated the case $P = Const$.

In order to control P, let us assume that the result of the measurement, i.e., a unit vector $a_m(t) = \{00\ldots0\underset{\underset{i}{\uparrow}}{1}0\ldots0\}$ is combined with an arbitrary complex (interference) vector.

If the interference state vector is

$$a' = \begin{pmatrix} a'_0 \\ a'_1 \\ \vdots \\ a'_N \end{pmatrix} \tag{10}$$

and σ is a measurement operator in the computational basis, then $|\psi(t+\tau)\rangle$, the recurrent state re-entering the circuit, must take one of the forms:

$$|\phi_0\rangle = \frac{1}{\sqrt{R_0}}\begin{pmatrix}1+a_0' \\ a_1' \\ \vdots \\ a_{N-1}'\end{pmatrix} = \frac{1}{\sqrt{R_0}}\begin{pmatrix}a_0^{(0)} \\ a_1^{(0)} \\ \vdots \\ a_{N-1}^{(0)}\end{pmatrix}$$

$$|\phi_1\rangle = \frac{1}{\sqrt{R_1}}\begin{pmatrix}a_0' \\ 1+a_1' \\ \vdots \\ a_{N-1}'\end{pmatrix} = \frac{1}{\sqrt{R_1}}\begin{pmatrix}a_0^{(1)} \\ a_1^{(1)} \\ \vdots \\ a_{N-1}^{(1)}\end{pmatrix} \tag{11}$$

$$|\phi_{N-1}\rangle = \frac{1}{\sqrt{R_{N-1}}}\begin{pmatrix}a_0' \\ a_1' \\ \vdots \\ 1+a_{N-1}'\end{pmatrix} = \frac{1}{\sqrt{R_{N-1}}}\begin{pmatrix}a_0^{(N-1)} \\ a_1^{(N-1)} \\ \vdots \\ a_{N-1}^{(N-1)}\end{pmatrix}$$

with re-normalization factors:

$$R_0 = |1+a_0'|^2 + |a_1'|^2 + \ldots \tag{12}$$

$$R_1 = |a_0'|^2 + |1+a_1'|^2 + \ldots \tag{13}$$

$$\vdots$$

$$R_{N-1} = |a_0'|^2 + |a_1'|^2 \ldots + |1+a_{N-1}'|^2 \tag{14}$$

It should be emphasized that the states (11) are first calculated and then prepared as new quantum inputs.

The transition probability matrix, p_{ij} for this process is given by examining how each of the recurrent states, $|\phi_0\rangle \cdots |\phi_{N-1}\rangle$ evolve under the action of U:

$$p_{ij} = \begin{pmatrix} \left|\frac{b_0^{(0)}}{\sqrt{R_o}}\right|^2 & \left|\frac{b_1^{(0)}}{\sqrt{R_0}}\right|^2 & \cdots \\ \left|\frac{b_0^{(1)}}{\sqrt{R_1}}\right|^2 & \left|\frac{b_1^{(1)}}{\sqrt{R_1}}\right|^2 & \cdots \\ \vdots & \vdots & \ddots \\ \left|\frac{b_0^{(N-1)}}{\sqrt{R_{N-1}}}\right|^2 & \cdots & \left|\frac{b_{N-1}^{(N-1)}}{\sqrt{R_{N-1}}}\right|^2 \end{pmatrix} \tag{15}$$

where

$$b_j^{(i)} = \sum_{\ell=0}^{N-1} u_{j\ell} a_\ell^{(i)} = u_{ji} + \sum_{\ell=0}^{N-1} u_{j\ell} a_\ell(0) \tag{16}$$

Thus, now the structure of the transition probability matrix p_{ij} can be controlled by the interference vector (10), and $P = P(t)$.

Let us now implement the internal feedback (4). For that purpose, assume that the components of the interference vector (10) are defined by the components π_i of the probability vector by setting:

$$a_i' = f_i(\pi_1, \pi_2, \ldots \pi_N) \tag{17}$$

and rewriting Eqs. (12) - (16) accordingly. Then

$$p_{ij} = p_{ij}(\pi_1, \ldots \pi_N) \tag{18}$$

However, the simplicity of this mathematical operation is illusive. Indeed, as pointed out above, the probability vector π is not simulated by the QRN explicitly: it has to be reconstructed by a statistical analysis of the ensemble of solutions to Eq. (5). In order to avoid that, one can simulate the evolution of the probability vector, i.e., Eq. (1) by a classical neural network which can be presented, for instance, in the form

$$\pi_i(t+\tau) = S\left[\sum_{j=1}^{N} w_{jk} \pi_k(t)\right] \tag{19}$$

where S is the sigmoid function, and $w_{jk} = Const$ are the synaptic weights.

Now Eqs. (5) and (19) are coupled via the feedbacks (6) and (17).

From the mathematical viewpoint, this system can be compared with the Langevin equation which is coupled with the corresponding Fokker-Planck equation such that the stochastic force is fully defined by the current probability distributions, while the diffusion coefficient is fully defined by the stochastic force.[4]

From the physical viewpoint, Eqs. (5) and (19) represent two different physical systems (quantum and classical) which interact via the feedbacks (4) and (6): the transition probability matrix P is defined by the unitary matrix U of the QRN according to Eq. (6), while the input interference vector to the QRN is defined by the feedback (17). Using the Feynmann terminology [1], Eq. (5) simulates probabilities, while Eq. (19) manipulates by them.

Finally, from the cognitive viewpoint, Eqs. (5) and (19) represent two different aspects of the same subject: the decision maker. Eq. (5) simulates his real-time actions, i.e., his motor dynamics, while Eq. (19) describes evolution of self-image in terms of such invariants as expectation, variance, entropy (information), and that can be associated with the mental dynamics.

Thus, as a result of interaction with his own image and without any "external" enforcement, the decision maker can depart from the thermodynamical limit (3) of his performance "against the second law." Obviously, from the physical viewpoint, the enforcement in the form of the feedback (17) is external since the image (19) represents a different physical system. In other words, such a "free will" effort is not in a disagreement with the second law of thermodynamics.

Eqs. (5) and (19) illuminate another remarkable property of human activity: the ability to predict future. Indeed, Eq. (19) depends only upon the prescribed unitary matrix

U, but it does not depend upon the evolution of the vector a_i. Therefore, Eq. (19) can be run faster than real time; as a result of that, future probability distributions as well as its invariants can be predicted and compared with the objective. Based upon this comparison, the feedback (17) can be changed if needed.

Actually such interaction with self-image simulates "common sense" which replaces an unavailable external source of information and allows one to make decisions based upon his previous experience.

Formally the knowledge base is represented by the synaptic weights w_{jk} of Eq. (19), and it consists of two parts. The first part includes personal experience and habits (risk prone, risk aversion, etc.). The second part depends upon the objective formulated in terms of probability invariants (certain expectation with minimal variance, or maximum information, etc.). The dependence upon the objective may include real-time adjustment of synaptic weights w_{ij} in the form of learning (adapted from theory of neural networks). As soon as the synaptic weights are determined, the common sense simulator will follow the optimal strategy regardless of unexpected changes in the external world.

It should be noticed that the advantage of the quantum implementation is not only in simulation of true randomness, but also in exponential increase of information capacity. Indeed, combining the direct product decomposability and entanglement,, one can represent the unitary matrix in Eq. (5) as follows:

$$U = \left(U_1^{(1)} \otimes \cdots U_n^{(1)}\right) \bullet \left(U_1^{(2)} \otimes \cdots U_n^{(2)}\right) \cdots \left(U_1^{(m)} \otimes \cdots U_n^{(m)}\right) \cdots \tag{20}$$

Here the number of independent components is:

$$q = 4nm \tag{21}$$

while the dimensionality

$$N = 2^n = 2^{q/4m} \tag{22}$$

In Eq. (22), N and q are associated with the Shannon and the algorithmic complexity, respectively; therefore, the exponential Shannon complexity is achieved by linear resources.

Further compression of Shannon information can be obtained by applying the $\ell-$measurement architecture [3] when each step of the quantum evolution is repeated and measured ℓ times, and during a reset operation the results of all the measurements are combined with the previous state. As shown in [3], such an architecture provides the double-exponential Shannon complexity:

$$N = 2^{q\ell/4m} \tag{23}$$

The advantage of the quantum compressions (22) or (23) can be appreciated in view of the fact that the efficiency of an alternative device - the pseudorandom number generator - rapidly decreases with the growth of the dimensionality of random vectors.

Finally, one should notice that QRN provides the simplest physical simulation of the four constraints in Eq. (1). However, even if QRN is replaced by a random number generator, the quantum formalism should be preserved since it is the best mathematical tool for implementation of these constraints.

2. Spontaneious self-organization

We will start the analysis of the motor-mental dynamics, i.e., of Eqs. (5) and (19) with the effect of a spontaneous self-organization when the system departs from the state of the thermodynamics limit and approches a deterministic state without any external forces. For that purpose suppose that the selected unitary matrix in Eq. (5) is

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{24}$$

Then the corresponding transition probability matrix in Eq. (1), according to Eq. (6) will be doubly-stochastic:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{25}$$

and the stochastic process (1) is already in its thermodynamics limit (3), i.e., $\pi_1 = \pi_2 = \frac{1}{2}$

Let us assume that the objective of the decision-maker is to approach the deterministic state

$$\pi_1 = 1, \quad \pi_2 = 0 \tag{26}$$

without help from outside. In order to do that, he should turn to his experience in the form of the feedback (17). If he chooses this feedback in the form:

$$a = (a_1, a_2), \; a_1 = -2\pi_1, \; a_2 = 1 \tag{27}$$

then, according to Eqs. (11-16), the new transition probability matrix p_{ij} will be:

$$\begin{aligned} p_{11} &= \frac{\pi_1^2}{2\pi_1^2 - 2\pi_1 + 1}, \quad & p_{12} &= \frac{(1-\pi_1)^2}{2\pi_1^2 - 2\pi_1 + 1} \\ p_{21} &= \frac{(\pi_1 + 1)^2}{2\pi_1^2 + 2}, \quad & p_{22} &= \frac{(1-\pi_1)^2}{2\pi_1^2 + 2} \end{aligned} \tag{28}$$

Hence, the evolution of the probability π_1 now can be presented as:

$$\pi_1^{(n+1)} = \pi_1^{(n)} p_{11} + \left(1 - \pi^{(n)}\right) p_{21} \tag{29}$$

in which p_{11} and p_{22} are substituted from Eqs. (28).

It is easily verifiable that

$$\pi_1^{\infty} = 1, \; \pi_2^{\infty} = 0 \tag{30}$$

i.e., the objective is achieved due to the "internal" feedback (27).

3. Attraction to common sense based strategies.

Classical artificial intelligence as well as artificial neural networks are effective in a deterministic and repetitive world, but faced with uncertainties and unpredictabilities, both of them fail. At the same time, many natural and social phenomena exhibit some degree of regularity only on a higher level of abstraction, i.e., in terms of some invariants. For instance, each particular realization of a stochastic process can be unpredictable in details, but the whole ensemble of these realizations i.e., "the big picture" preserves the probability invariants (expectation, moments, information, etc), and therefore, predictable in terms of behavior "in general."

In this section we will map the hetero-associative memory problem performed by artificial neural nets onto the patterns which represent stochastic processes, namely: store a set of m stochastic processes given by vectors of their probability distributions

$$\pi^{(i)} = \pi_1^{(i)}, \pi_2^{(i)} \ldots \pi_N^{(i)}, \quad i = 1, 2, \ldots m \tag{31}$$

in such a way that when presented with any of the process $\overset{*}{\pi}^{(j)}$ out of the set of M processes:

$$\overset{*}{\pi}^{(j)} = \overset{*}{\pi}_1{}^{(j)}, \overset{*}{\pi}_2{}^{(j)}, \ldots \overset{*}{\pi}_N{}^{(j)}, \; j = 1, 2, \ldots M, \tag{32}$$

the coupled motor-mental dynamics (5), (19) converges to one of the stochastic processes (31).

The performance

$$\overset{*}{\pi}^{(i)} \to \overset{*}{\pi}^{(i)}, \quad i = 1, 2, \ldots m; \tag{33}$$

represents correspondence between two classes of patterns, i.e., a hetero-associative memory on a high level of abstraction. Indeed, each process in (33) stores an infinite number of different patterns of behaviors which, however, are characterized by the same sequence of invariants (31) and (32), respectively thereby representing a decision making strategy.

Hence, if the strategy of the decision-maker is characterized by a pattern $\overset{*}{\pi}^{(i)}$ from (32), and starting from t=0, the external information becomes unavailable, he should change its strategy from the pattern $\overset{*}{\pi}^{(i)}$ to the corresponding pattern from (31), and that can be associated with a decision based upon common sense. It is implied that the attracting strageties $\pi^{(i)}$ are sufficiently "safe", i.e., they minimize the risk taken by the decision-maker in case of an uncertain external world.

The first step in the implementation of the mapping (33) is to find the transition probability matrix P such that

$$\pi^{(i)} = \overset{*}{\pi}^{(i)} P\left(\overset{*}{\pi}^{(1)}, \overset{*}{\pi}^{(2)} \ldots \overset{*}{\pi}^{(m)} \right) \tag{34}$$

This implies that the sought stochastic process is supposed to approach its limit state in one step, i.e.,

$$\overset{*}{\pi}^{(i)}(t+\tau) = \overset{*}{\pi}^{(i)}(t+2\tau) = \overset{*}{\pi}^{(i)}(\infty) = \pi^{(i)} \tag{35}$$

Therefore, P must have the following form:

$$P = \begin{pmatrix} \pi_1 & \dots & \pi_N \\ \pi_1 & \dots & \pi_N \\ \vdots & & \vdots \\ \pi_1 & \dots & \pi_N \end{pmatrix}, \quad 0 < \pi_i < 1, \quad \sum_{i=1}^{N} \pi_i = 1 \tag{36}$$

where the vector $\pi = (\pi_1, \dots \pi_N)$ belongs to the family of the vectors $\pi^{(i)}$ in Eq. (35).

Indeed, then any arbitrary probability vector

$$X = (x_1, x_2, \dots x_N), \quad \sum_i x_i = 1 \tag{37}$$

is mapped onto the vector $\pi = (\pi_1, \dots \pi_N)$ in one step.

Let us assume that the vector $\pi = (\pi_1, \dots \pi_N)$ is representable as a direct product of n two-dimdensional vectors.

$$(\pi_1, \pi_2 \dots \pi_N) \rightarrow (\pi_1 \;\; 1 - \pi_1) \otimes \dots \otimes (\pi_n, \; 1 - \pi_n) \tag{38}$$

$$n = \ell og_2 N \tag{39}$$

Obviously this assumption imposes constraints upon the components of the vector π, and as a result, this vector can be defined only by $\ell og_2 N$ (out of N) independent parameters $\pi_j, \quad i = 1, 2, \dots n$.

Now Eq. (36) reduces to

$$P = \begin{pmatrix} \pi_1 & 1 - \pi_1 \\ \pi_1 & 1 - \pi_1 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} \pi_n & 1 - \pi_n \\ \pi_n & 1 - \pi_n \end{pmatrix} \tag{40}$$

where

$$p_{11}^{(k)} = p_{21}^{(k)} = \pi_k, \qquad p_{12}^{(k)} = p_{22}^{(k)} = 1 - \pi_k$$

The next step in the implementation of the mapping (33) is to express the components of the matrix (40) via the components of the unitary operator U_{ij} (see Eq. (5)) and the interference vector (10). For that purpose, let us choose U_{ij} *and* α' as follows:

$$U = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{41}$$

$$a' = \left(a_1, a_{1(1)} + i\beta_{1(1)}\right) \otimes \ldots \otimes \left(a_n, a_{1(n)} + i\beta_{1(n)}\right) \tag{42}$$

Then, according to Eqs. (11)-(16),

$$p_{11}^{k} = \frac{|a_k + 1|}{|a_k + 1|^2 + \left|a_{k(1)}^2 + b_{k(1)}^2\right|} = \pi_k = p_{21}^{(k)} = \frac{|a_k|^2}{|a_k|^2 + \left|a_{k(1)} + b_{k(1)} + 1\right|^2} \quad k = 1,2,\ldots n \tag{43}$$

However, the components of the interference vector, $\alpha_k, \alpha_{k(1)}$ and $\beta_{k(1)}$ cannot be chosen independently since they should explore the equality (43) as well as the conditions:

$$I_m a_k = 0, I_m a_{k(1)} = 0,\ I_m b_{k(1)} = 0 \tag{44}$$

Simple algebra leads to the following constraints imposed upon the interference vector:

$$a_k > -1, \quad k = 1,2,\ldots n \tag{45}$$

$$a_{k(1)} = \frac{a_k^4}{2(a_k + 1)^2} - \frac{\left(a_k^2 + 1\right)}{2} \tag{46}$$

$$\beta_{k(1)} = \sqrt{a_k^2 - a_{k(1)}^2} \tag{47}$$

Now the components π_k in Eq. (43) can be expressed via the only one component of the interference vector:

$$\pi_k = \frac{(a_k + 1)^2}{(a_k + 1)^2 + a_k^2}, \quad 1 - \pi_k = \frac{a_k^2}{(a_k + 1)^2 + a_k^2} = \tilde{\pi}_k \tag{48}$$

It is easily verifiable that $\tilde{\pi}_k$ is a sigmoid function of a_k:

$$\tilde{\pi}_k = S(a_k) \ \text{since} \frac{\partial \tilde{\pi}_k}{\partial a_k} \geq 0, \quad \tilde{\pi}_k(0) = 0;\ \tilde{\pi}_k(\infty) = \frac{1}{2} \tag{49}$$

and that property will be exploited later.

The final step is to implement the actual association between the patterns in the mapping (33), i.e., to find the appropriate dependence between the components π_k of the matrix (40) and the components of the pattern $\overset{*}{\pi}^{(i)}$. Since π_k are uniquely defined by a_k (see Eqs. (48)), we will start with representing a_k as linear combinations of the components of the initial patterns $\overset{*}{\pi}^{(j)}$ in the mappint (33) for each j^{th} association:

$$a_k^{(j)} = \sum_{i=1}^{N} w_{ik} \overset{*}{\pi}_i^{(j)},\ j = 1,2\ldots m;\ k = 1,2,\ldots n \tag{50}$$

where w_{ik} are constant weights to be found, m is the number of associations in Eq. (33), N and n are the dimensionalities of the input pattern $\overset{*}{\pi}^{(i)}$ and the output pattern $\pi^{(j)}$, respectively.

Eq. (50) can be written in the matrix form

$$A_{mn} = W_{nN}\, \Pi_{mN} \tag{51}$$

and therefore, the matrix W_{nN} of the weights can be explicitely expressed via the matrix A_{mn}, i.e., via the components of the interference vector $a_k^{(j)}$:

$$W_{nN} = \begin{cases} A_{mn}\Pi_{NN}^{-1} \text{ if } m = N, \det \Pi \neq 0 & (52) \\ A_{mn}\left(\Pi^T\Pi\right)^{-1}\Pi^T \text{ if } m > N & (53) \end{cases}$$

Eq. (52) presents the exact solution, while Eq. (53) gives a minimum norm approximateion for the case when the number of association is larger than the dimensionality of the input patterns $\overset{*}{\pi}^{(j)}$.

Since $a_k^{(j)}$ can be expressed via the probabilities $\pi_k^{(j)}$ of the transition probability matrix (38) by means of Eq. (48):

$$a_l^{(j)} = \frac{2\tilde{\pi}_k^{(j)} \pm \sqrt{12\left(\tilde{\pi}_k^j\right)^2 - 4\tilde{\pi}_k^{(j)}}}{2\left(1 - 2\tilde{\pi}_k^{(j)}\right)} \tag{54}$$

(one can choose either of two values), the problem is solved in a closed analytical form. Indeed, given the associations (33), one finds the corresponding $a_k^{(j)}$ by Eqs. (54), and then the weights w_{ij} depend upon all the values of the input patterns $\overset{*}{\pi}_k^{(j)}$ (via the matrix Π) and the output patterns $\pi_k^{(j)}$ (via the matrix A).

As soon as the weights w_{ij} are found, Eq. (19) can be represented in the following form:

$$\pi_i^{\infty} = S\left(\sum_{i=1}^{N} w_{ik}\pi_k^{o}\right), \; i = 1, 2, \ldots N \tag{55}$$

where $$\pi_i^{\infty} = \pi_i(t \to \infty), \quad \pi_k^{o} = \pi(t = 0) \tag{56}$$

and the sigmoid function S is defined by Eq. (49).

Eq. (55) has a form of a perceptron for hetero-associative memory. Exploiting this formal analogy, one can conclude that any input pattern π^{o} which is sufficiently close to a pattern $\pi^{(i)}$ from the left of Eq. (33) will recall the output pattern which is close to the corresponding associative pattern $\overset{*}{\pi}^{(i)}$ from the right of Eq. (33). Moreover, due to the contracting property of the sigmoid function S in Eq. (55), the distance between the output patterns will be smaller than between the input ones. In particular, several different inputs can be mapped onto the same output, and that can be interpreted as a classification problem.

However, from the cognitive viewpoint, Eq. (55) is fundamentally different from the perceptron since it not only manipulates with the patters of probabilities, but it also simulates them via the QRN. Indeed, Eqs. (50) defines the interference vector a' (see Eqs. (42)) which control the unitary evolution of QRN (see Eqs. (5) and Eq. (41)) in such a way that the generated stochastic process has exactly the same probability distribution as prescribed by the probability pattern π^{∞} manipulated by Eq. (55).

Thus we have introduced a new dynamical paradigm in the form of coupled motor and mental dynamics which is represented by a quantum generator of stochastic processes controlled by nonlinear Markov chains. Based upon this paradigm, a quantum decision-maker has been proposed. New dynamical phenomena, namely spontaneous self-organization, and attraction to common sense strategies have been discussed.

Acknowledgments

The research described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology and was sponsored by the National Aeronautics and Space Administration, Office of Space Science and Technology.

1. R. Feynmann, Int. J. of Theoretical Physics, Vol. 21. No. 6/7, 1982.

2. E. Schrödinger, What is Life, Cambridge University Press, Cambridge, 1944.

3. M. Zak, C. Williams, Quantum Neural Nets, Int. J. of Theor. Physics, Feb. 1998.

4. M. Zak, Physical Invariants of Biosignatures, Physics Letters A, 1999